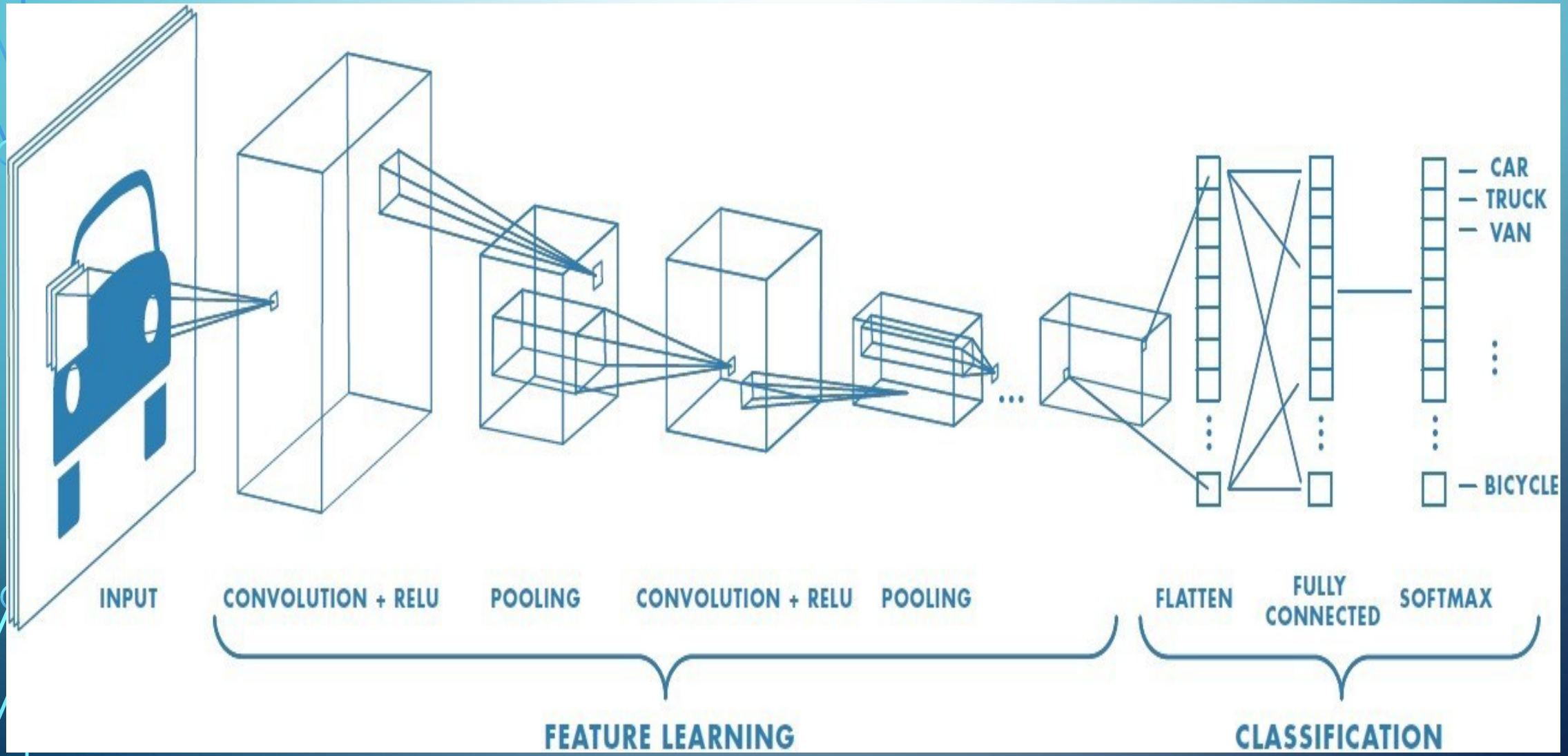
CONVOLUTIONAL NEURAL NETWORKS(CNN) – NOT THAT CNN ᴍ

1. SHORT INTRODUCTION TO WHAT CONVOLUTIONAL NEURAL NETWORKS ARE
2. HOW CNN WORK
3. PATTERNS WHEN DESIGNING CNNS
4. KNOWN CNN ARCHITECTURES(ALEXNET, VGGNET, GOOGLENET, RESNET)
5. SOME CODE AND SOME FUN USES ᴍ

CNNs work in the same way as regular ANN, each layer of a CNN finds succesively complex features

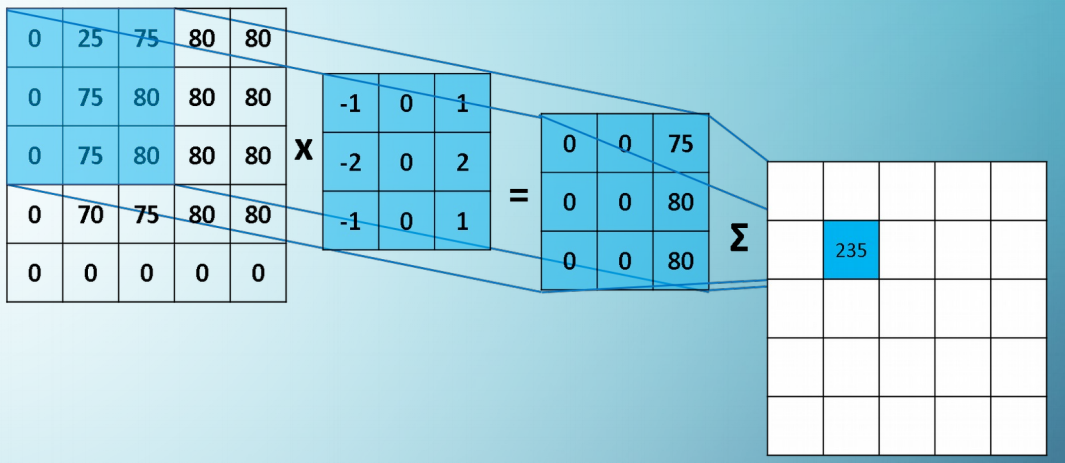
- They stack neurons in hidden layers
- They have weights that are randomly initiated and learned durin training
- They still apply an activation function
- They still calculate the error and then backpropagte the error to update the weights
- The main difference is that they use ”special” layers/modules called conolution layers instead of fully connected layers

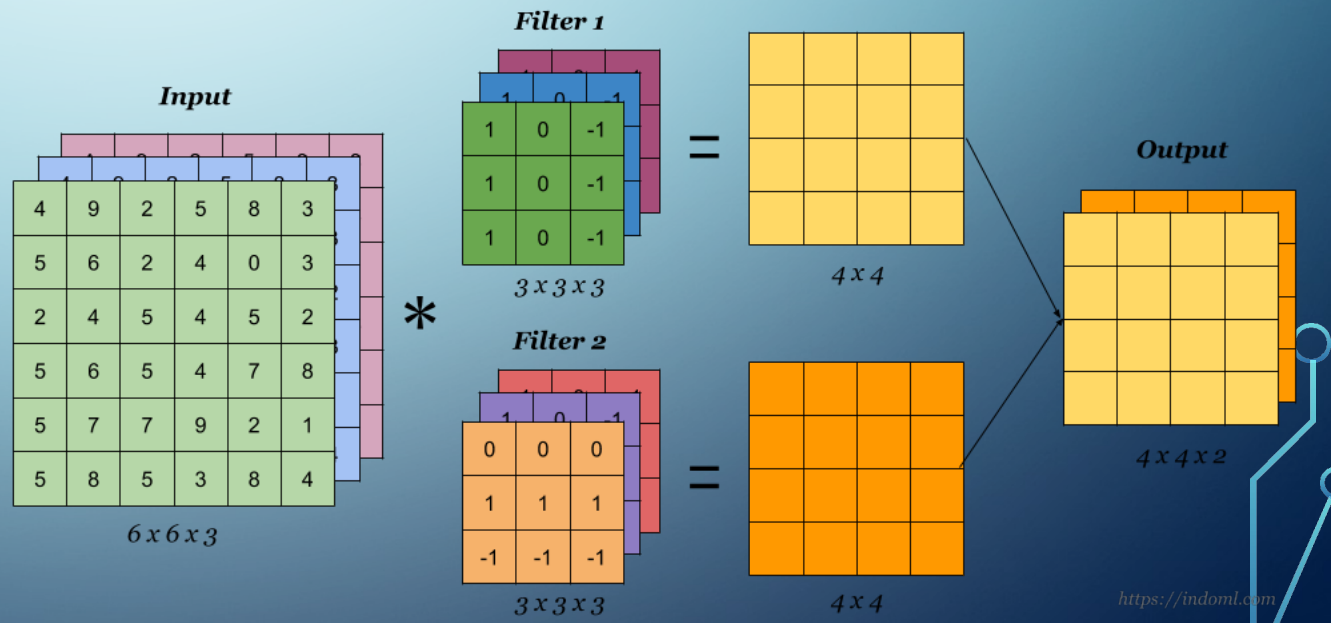

CAR
TRUCK
VAN
BICYCLE
INPUT
CONVOLUTION + RELU
POOLING
CONVOLUTION + RELU
POOLING
FLATTEN
FULLY CONNECTED
SOFTMAX
FEATURE LEARNING
CLASSIFICATION

WHY ARE CONVOLUTIONS BETTER THAT FULLY CONNECTED NETWORKS?

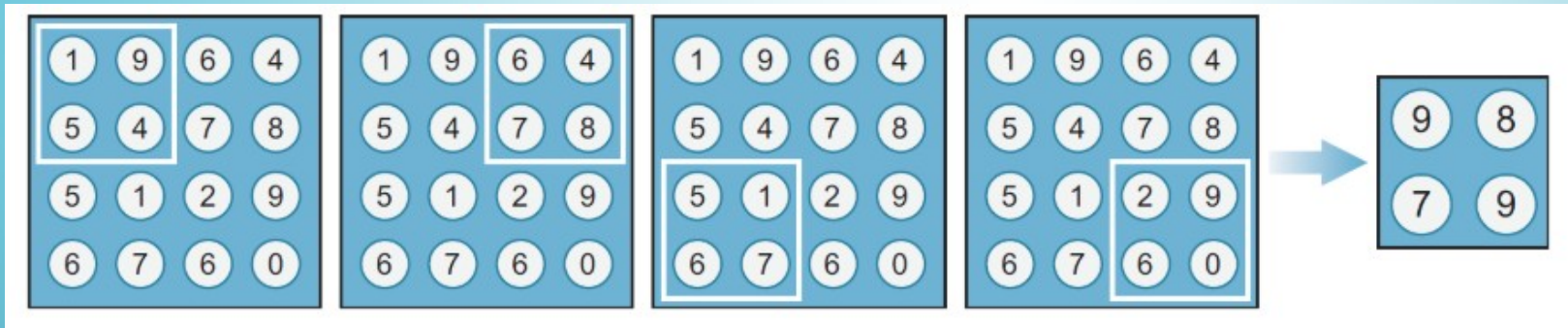
- MLPs have to flatten the image to be able to process it(loos of spatial features), while CNNs can treat raw images as the input data
- MLPs have fully connected layers that yield a very high computational cost, while CNN are locally connected(the nodes are connected to a small subset of the previous layers node)
- An image of 1000 X 1000 and 1000 neurons/hidden layer a MLP would do 1 bilion operations just for the first layer!

- Kernels to the rescue(a fancy name for 2D matrix, also known as filters)
- In CNNs the kernel contain the weights that the network needs to learn.

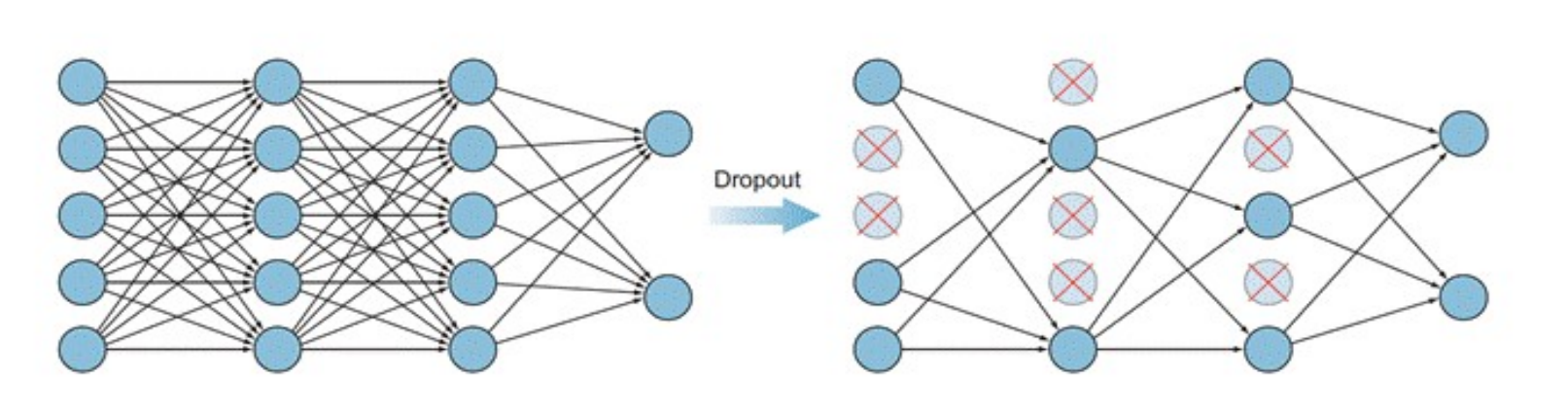




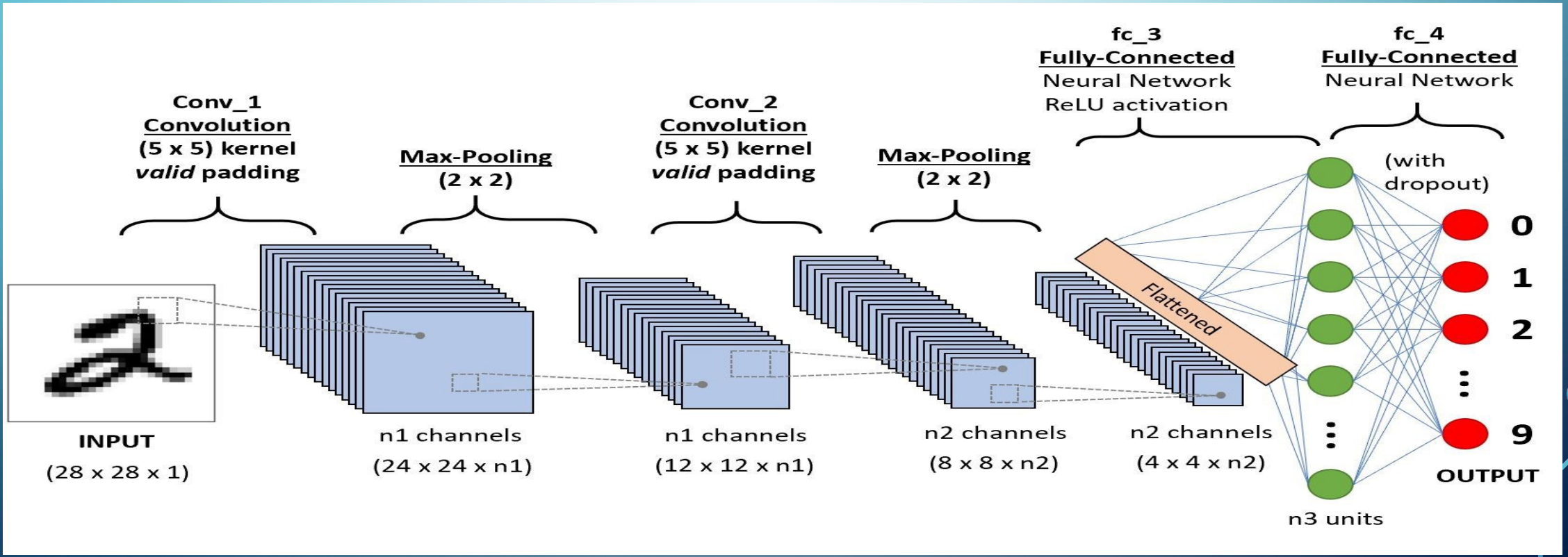
POOLING(SUBSAMPLING)



DROPOUT



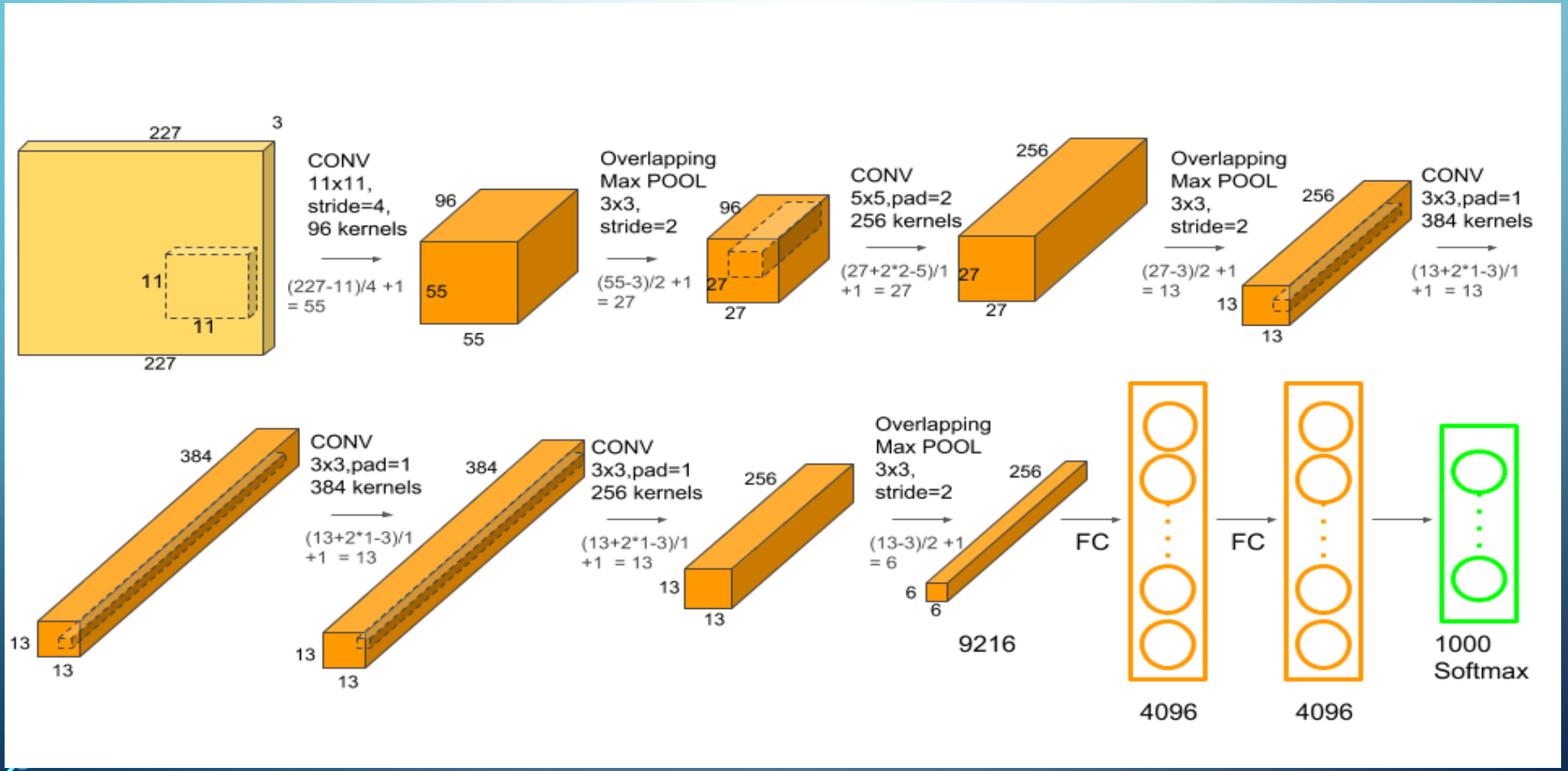
PUTTING IT ALL TOGHETER



PATTERNS (BECAUSE EVERYONE LOVES THEM)

1. CNNs have two parts: a feature extraction part which consists of a series of convolutional layers and a classification part that consists of a series of fully connected layers. Use RELU for hidden layers and SOFTMAX for classification layer
2. As the input (image) passes through the network depth is increasing and dimension is decreasing. This translates to that network is learning more complex features as the image passes through the network
3. Last part of the network are fully connected layers that decresease in size

ALEXNET


227
3
227
11
11
CONV
11x11,
stride=4,
96 kernels
(227-11)/4 +1
= 55
96
55
55
Overlapping
Max POOL
3x3,
stride=2
(55-3)/2 +1
= 27
96
27
27
CONV
5x5,pad=2
256 kernels
(27+2*2-5)/1
+1 = 27
256
27
27
Overlapping
Max POOL
3x3,
stride=2
(27-3)/2 +1
= 13
256
13
13
CONV
3x3,pad=1
384 kernels
(13+2*1-3)/1
+1 = 13
384
13
13
CONV
3x3,pad=1
384 kernels
(13+2*1-3)/1
+1 = 13
384
13
13
CONV
3x3,pad=1
256 kernels
(13+2*1-3)/1
+1 = 13
256
13
13
Overlapping
Max POOL
3x3,
stride=2
(13-3)/2 +1
= 6
256
6
6
9216
FC
4096
FC
4096
1000
Softmax

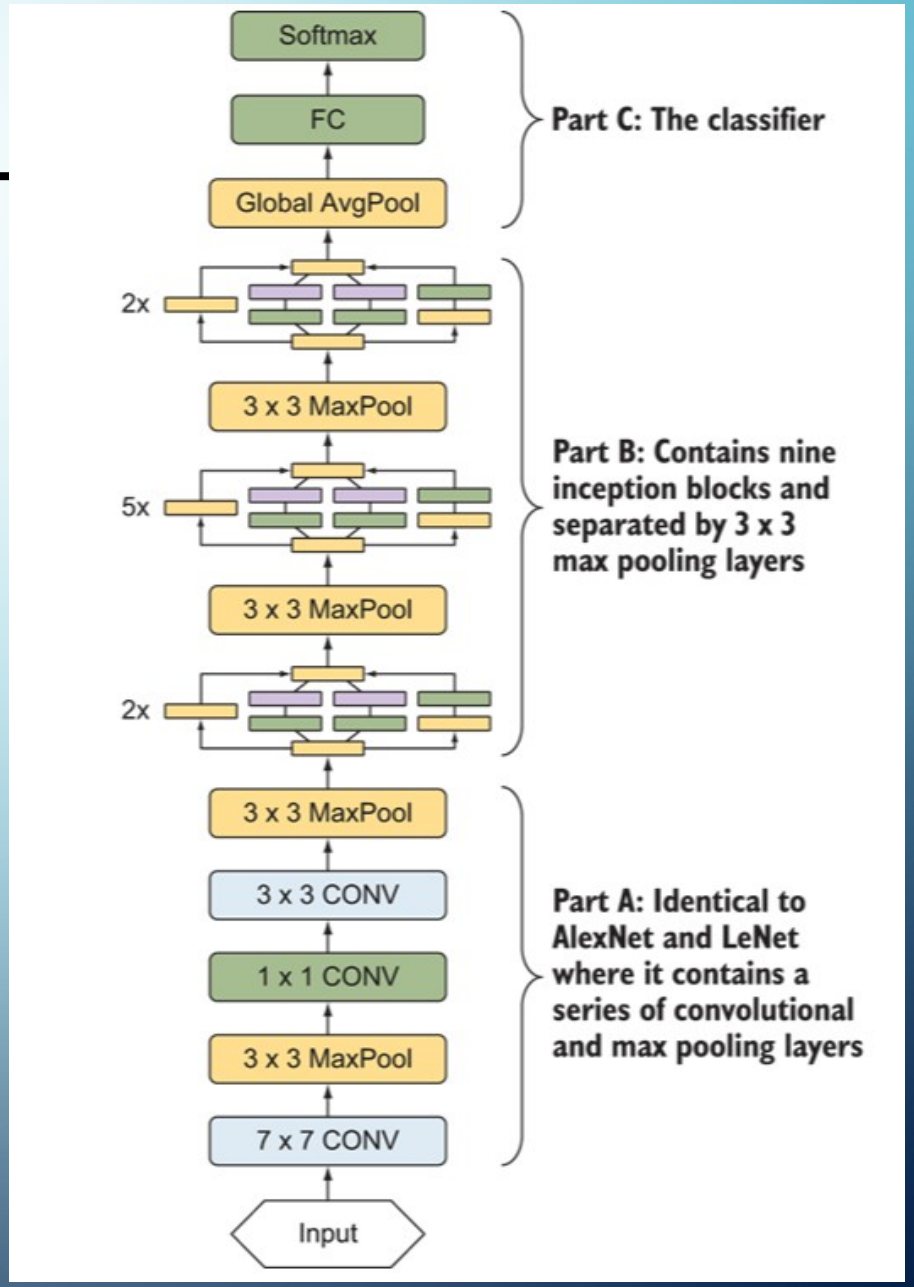
- Activation function: RELU(to avoid vanishing gradient)
- Dropout: used only in fully connected layers after each one
- Data Augmentation: the images was flipped and rotated in different positions
- Batch Normalization to avoid Covariate Shift
- L2 regularization
- Max Pooling

VGG - VISUAL GEOMETRY GROUP AT OXFORD UNIVERSITY

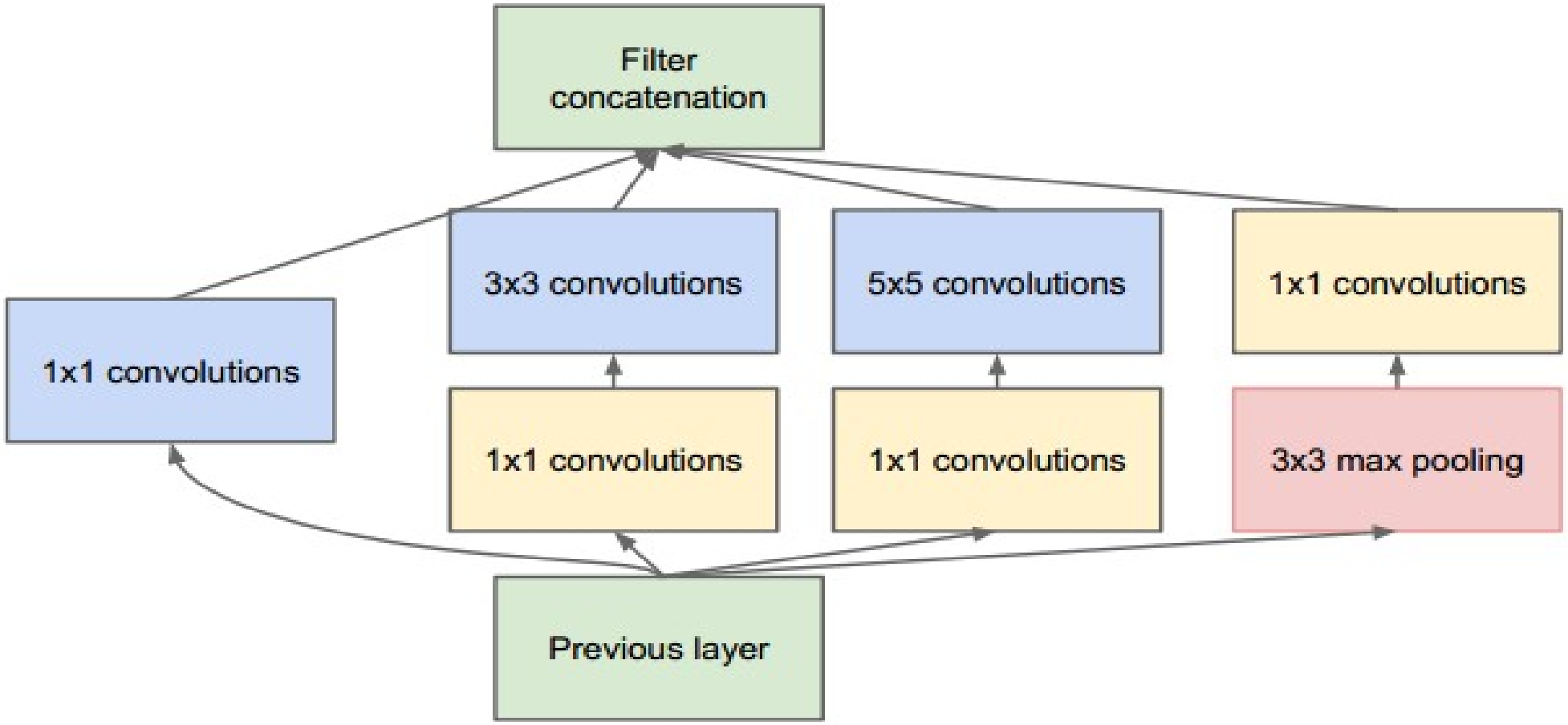
- No new „block” from AlexNet
- The mai difference is that in VGGNet they use 3X3 pool size filter and not of varying size
- Two 3X3 filters with no pooling between them act like a 5X5 filter, three act as a 7X7. Decreases the number of training params.

INCEPTION/ GOOGLENET

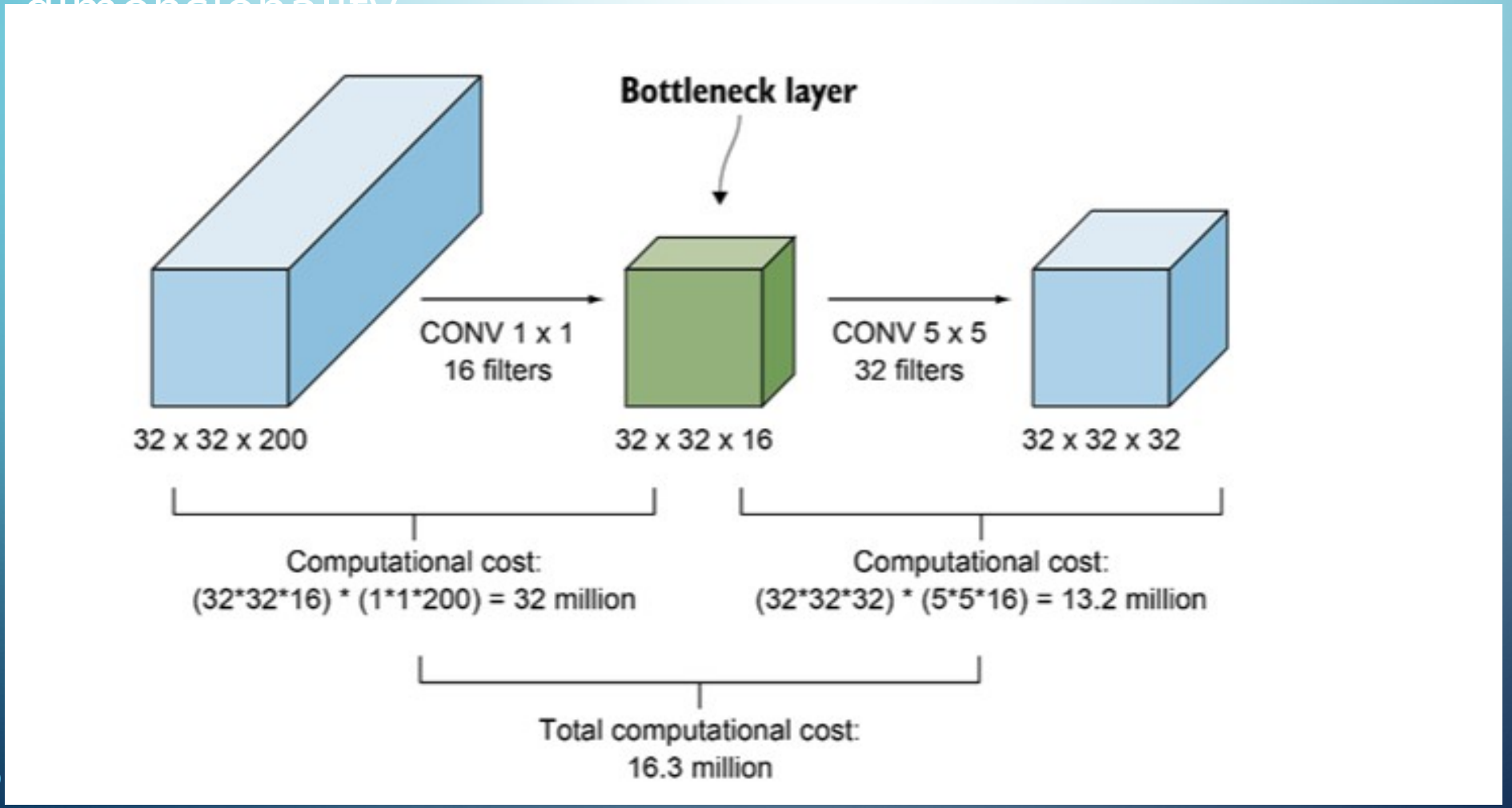
Is formed from multiple inception modules that contain various COV layers with different pool sizes



INCEPTION MODULE



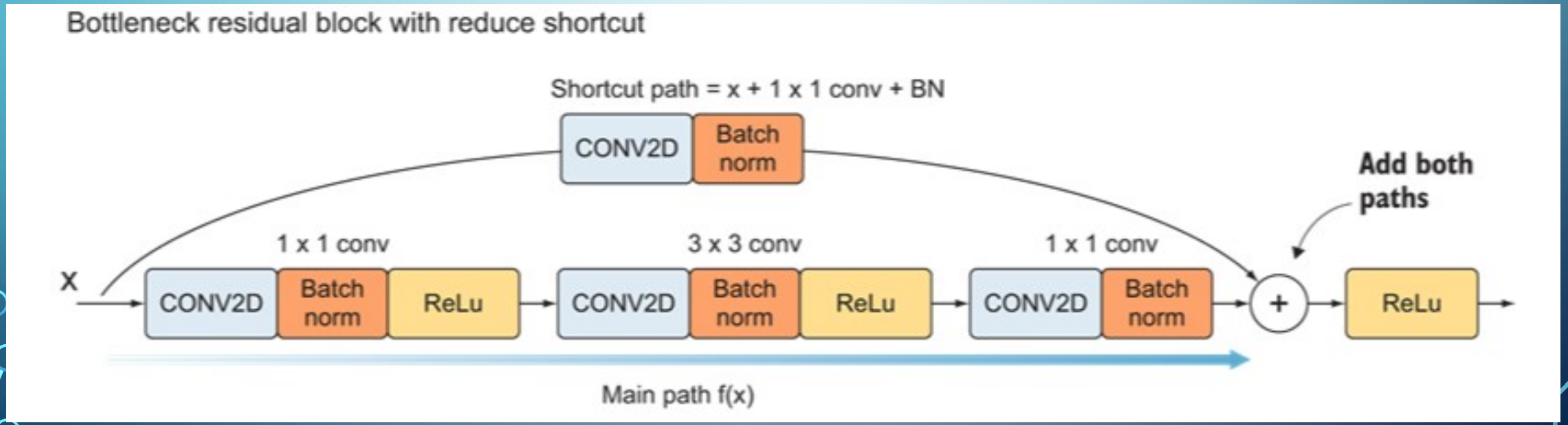
- The 1X1 componenet is just for depth reduction. It will maitain the height and width of the input but will reduce the dimensionality



RESNET

- Was the first network to test how deep can a neural network can be architected
- The nouvel features were the ability to bypass the vanishing and exploding gradient by using a ”skip connection”

RESIDUAL BLOCK



Year	CNN	Number of layers	Top-5 error rate	Number of parameters
1998	LeNet	5 layers	NA	60 thousand
2012	AlexNet	8 layers	15.3%	60 million
2014	VGGNet	16 layers	7.3%	138 million
2014	Inception (GoogLeNet)	22	6.67%	12 million
2015	ResNet-152	152	4.49%	